AFRL-RY-WP-TR-2011-1118

CENTER FOR AUTOMATIC TARGET RECOGNITION RESEARCH

Delivery Order 0005: Image Georegistration, Camera Calibration, and Dismount Categorization in Support of DEBU from Layered Sensing

Alper Yilmaz

The Ohio State University

JULY 2011
Final Report

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

*//Signature//

KELLY MILLER, Project Engineer
Layered Sensing Performance
Characterization Branch
Layered Sensing Exploitation Division

//Signature//

CHRISTINA SCHUTTE, Branch Chief
Layered Sensing Performance
Characterization Branch
Layered Sensing Exploitation Division

//Signature//

CHRIS RISTICH, Division Chief
Layered Sensing Exploitation Division
Sensors Directorate

*Disseminated copies will show "//Signature//" stamped or typed above the signature blocks.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
July 2011	Final	09 April 2010 – 01 April 2011

4. TITLE AND SUBTITLE
CENTER FOR AUTOMATIC TARGET RECOGNITION RESEARCH
Delivery Order 0005: Image Georegistration, Camera Calibration, and Dismount Categorization in Support of DEBU from Layered Sensing

5a. CONTRACT NUMBER
FA8650-07-D-1220-0005

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62204F

6. AUTHOR(S)
Alper Yilmaz

5d. PROJECT NUMBER
6095

5e. TASK NUMBER
19

5f. WORK UNIT NUMBER
6095190D

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Ohio State University
2070 Neil Avenue
Columbus, OH 43210

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RYAA

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RY-WP-TR-2011-1118

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
PAO Case Number: AFMC-2011-0214; Clearance Date: 18 Oct 2011. This report contains color.

The Appendix requires the use of 3-D glasses to achieve the intended full visual effect.

14. ABSTRACT
Aerial multi-head camera systems can provide a single synthetic image with large coverage from a set of images acquired simultaneously at particular time instant from each camera. To generate precise synthetic images, it is important to know the geometry between camera heads. In this project, DSM (Digital Surface Model) was generated from images acquired from aerial multi-head camera system. The process for generating DSM can be divided into two parts; Mosaic image generation and DSM generation. Relative position and orientation of physical cameras (six cameras) are given in terms of CAHVOR models. Mosaic images were generated projecting images acquired from physical cameras to a synthetic camera model via the reference plane. The synthetic camera for mosaic image is precisely designed to minimize resampling error. The EOP (exterior orientation parameters) of generated mosaic images can be calculated from navigation solutions which are given in pos file. 3D coordinates of a point can be calculated by space intersection of conjugate points of a pair of images with known IOP and EOP. Conjugate points can be automatically obtained by using image matching methods. Result of the image matching is 3D coordinates of a point. DSM can be generated by interpolating matching results.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	b. ABSTRACT	c. THIS PAGE			Kelly Miller
Unclassified	Unclassified	Unclassified	SAR	36	19b. TELEPHONE NUMBER *(Include Area Code)* N/A

Table of Contents

Section **Page**

List of Figures

Figure **Page**

1. Mosaic Generation

In this quarter we improve upon the algorithms developed in the last quarter to make them more robust and reliable. The software was improved for ease of operation and future extensibility. The initial estimation of homographies over the image overlap graph – a crucial step determining the accuracy of the final result – was improved. Up till now, a simple greedy algorithm traversing over the strongest connections in the graph was being used. This was replaced with a more robust global shortest path algorithm. The 'length' of an edge on the graph was defined to be 1 – strength where strength is the number of tied points linking the two images. This method gives better initial estimates for the homographies.

The datasets we are dealing with are typically large. The test dataset provided has about 1500 x 6 = 6000 high resolution images. To apply our algorithm, point matches need to be found between all pairs of images. It is easy to see that total number of matches to be evaluated grows rapidly.

We devised a user controllable scheme to deal with this data complexity. The user chooses a frame interval at which they wish to view the generated mosaic. Point matches are then evaluated only within that interval. The user can then 'flip' through the mosaics. Additionally, a small scale panorama over a large number of frames may also be generated. Since the data processing takes a lot of time, it is important that the data generation and caching be robust. We provide for a multi-stage processing of data which can be started and stopped at any stage. We also provide for a smart caching system such that data cached in one system can be used at any later stage at another system.

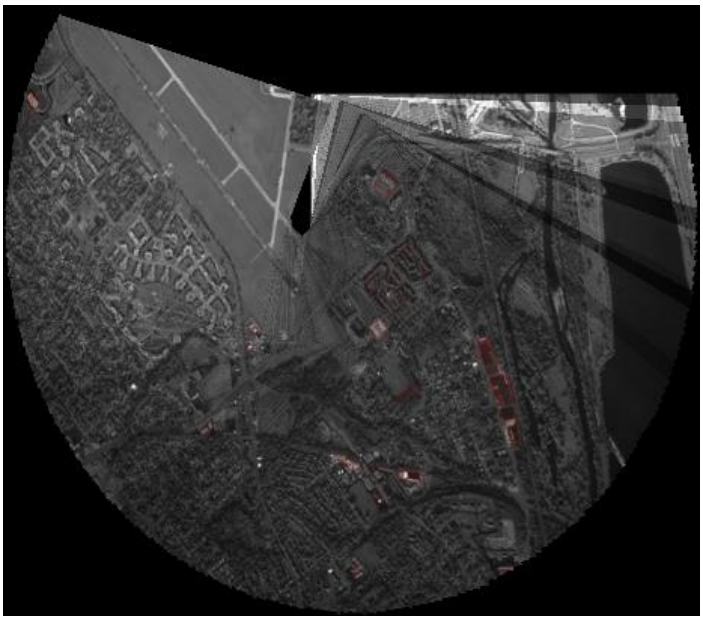


Figure 1. Mosaic of camera 0 over 200 frames. Potential occluders overlayed in red.

2. 3D Recovery from Images

The objective of this work is generating DSM (Digital Surface Model) from the aerial multi-head camera system. We have developed two software that generate synthetic mosaic images from raw images acquired from the multi-head camera system and 3D point cloud from the synthetic mosaic images. In this report, brief underlining principles and modification since final report are introduced.

2.1 Mosaic image generation

2.1.1 The synthetic camera model

To generate accurate 3D DSM, synthetic images should be generated very accurately. The synthetic perspective center and the synthetic focal length are precisely selected to minimize errors due to positional displacement of real perspective centers and under and over sampling. Figure 2 shows the generated synthetic camera (red) and given six physical cameras (blue).

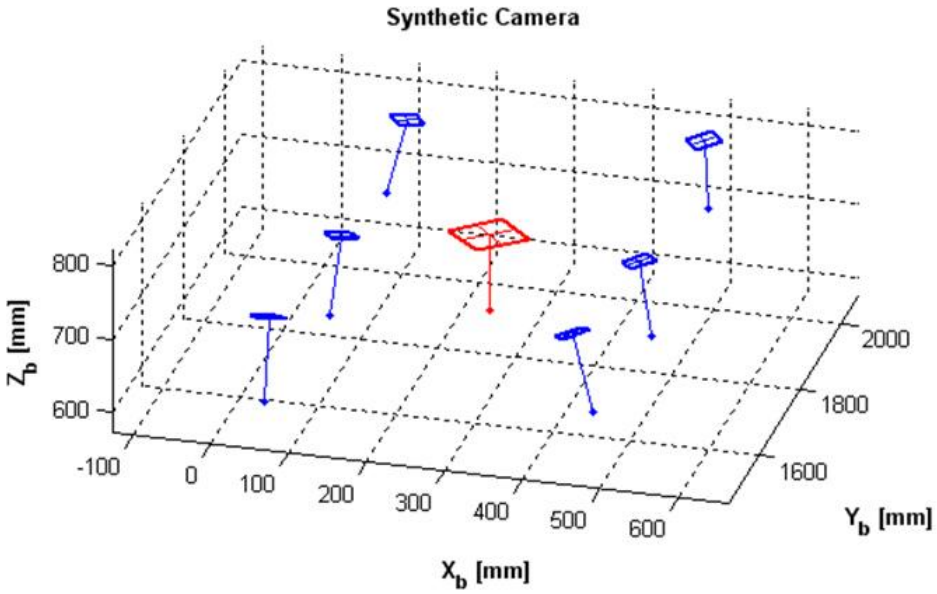


Figure 2. Six physical cameras (blue) and the synthetic camera for the mosaic image (red)

2.1.2 Mosaic image generation

Mosaic image can be generated by the approximation model that re-projects images from physical cameras via the reference plane (Figure 3). The error due to surface undulation is ignorable when ratio between surface undulation and flying height ($\Delta h/h_g$) is less than 0.2. The rigorous model needs true surface model of target area which is not always available.

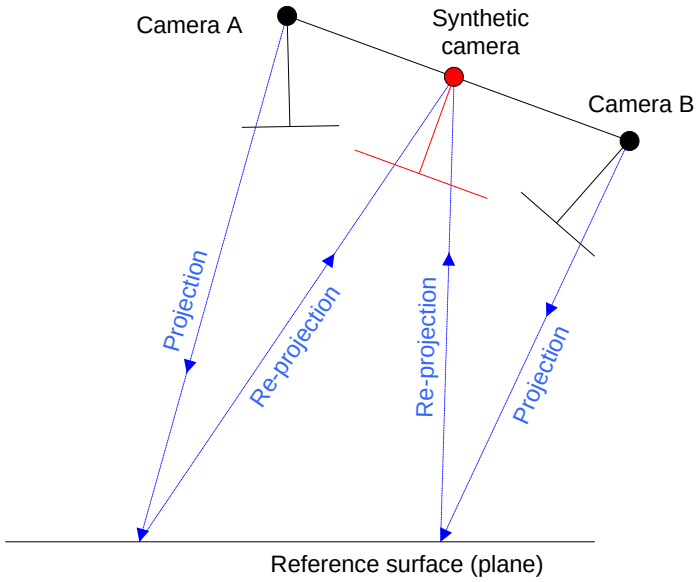


Figure 3. Conceptual diagram of mosaic image generation

The edges of mosaic images have poor geometric properties due to radial lens distortion as well as obliqueness of images. Previously, we generated mosaic images which include all pixels from raw images. As a result, mosaic image had null area (Figure 4; left) which is not included in raw images. We eliminate these null areas, which ill-affect image matching for 3D point generation and have poor geometric properties, by limiting the size of the mosaic image. Consequently, the size of the mosaic image is changed from (9742 by 10058) to (9400 by 9400).

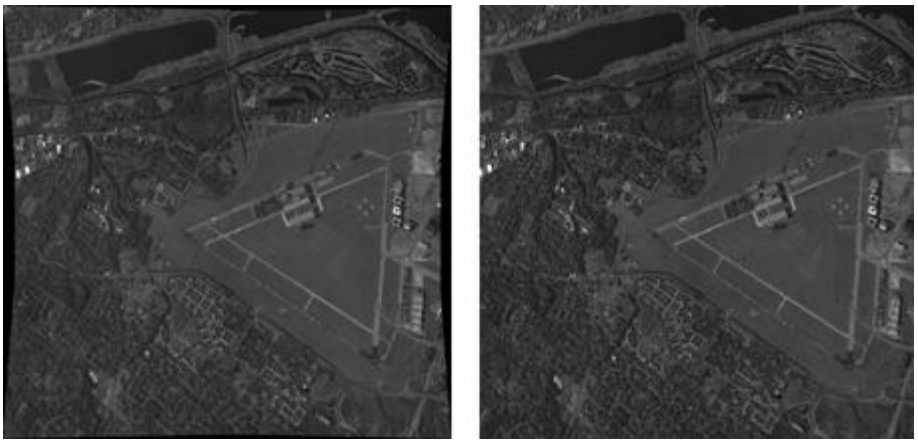


Figure 4. A mosaic image with null areas (left) and a mosaic image without null area (right)

2.1.3 Radiometric correction

The mean-standard deviation method is used for the radiometric correction. The method uses mean and standard deviation of pixels in overlapping areas between images. The method can be expressed as followings.

Let m_1, m_2 are means and s_1, s_2 are standard deviations of pixel values in overlapping area of the image 1 and 2, respectively. Then, radiometrically corrected pixel values of image 2 can be calculated by following equation.

$$y = ax + b;$$

$$a = \frac{s_1}{s_2},\ b = m_1 - am_2$$

where, y is new pixel value; x is old pixel value of image 2.

Exact overlapping areas between images are calculated from given CAHVOR model (previously, rough overlapping areas were used) for the radiometric correction. Figure 5 illustrates the result of the radiometric correction.

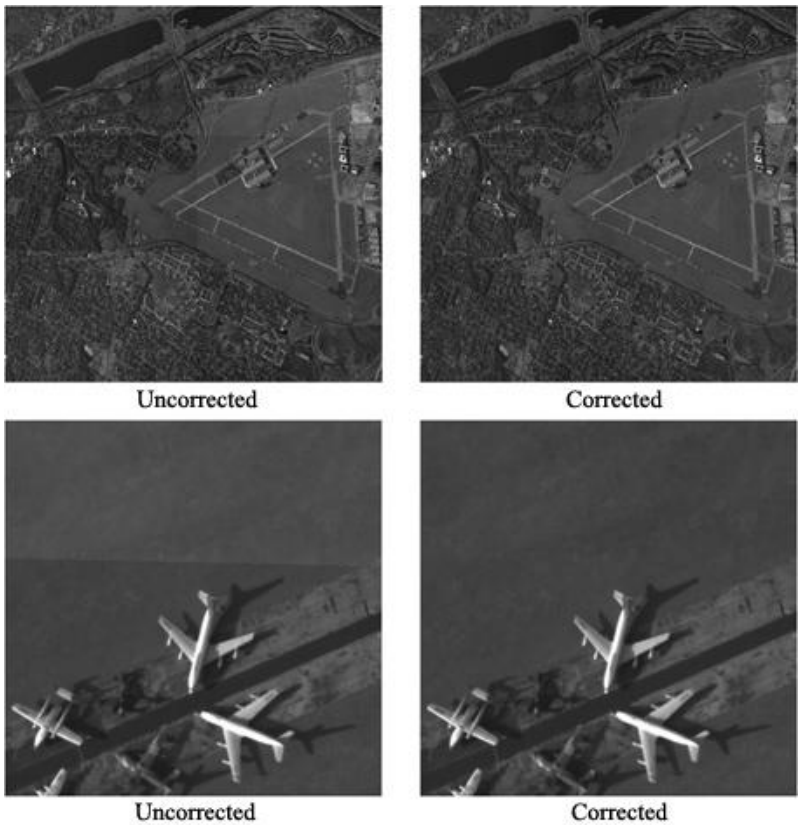


Figure 5. Mosaic images without radiometric correction (left) and with radiometric correction (right)

2.2 DSM generation

2.2.1 Image matching

We use correlation matching that calculates correlation coefficient between moving template (from image A) and search space (from image B). Figure 6 illustrates concept of the correlation matching in image space. We generate epipolar image pair of which rows of the images have same information.

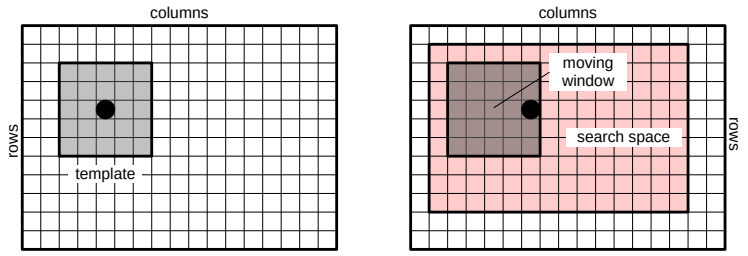


Figure 6. Concept of the correlation matching in image space

2.2.2 Epipolar line constraint

We use the epipolar line constraint that a point is correspond to a line in a stereo pair to reduce search space in row direction. Figure 7 shows concept of the epipolar line constraint in image space and Figure 9 (left) illustrates geometry of the epipolar constraint.

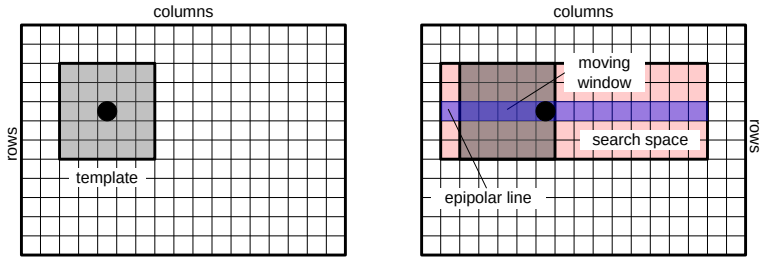


Figure 7. Concept of the epipolar line constraint in image space

2.2.3 Vertical line locus constraint

If the maximum and the minimum heights of target area are known, epipolar lines can be reduced. Therefore, search space is reduced in column direction. Figure 8 shows concept of the vertical line locus constraint in image space and Figure 9 (right) illustrates geometry of the vertical line locus constraint.

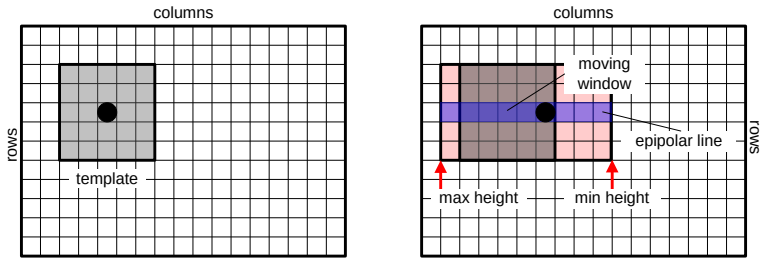


Figure 8. Concept of the vertical line locus constraint in image space

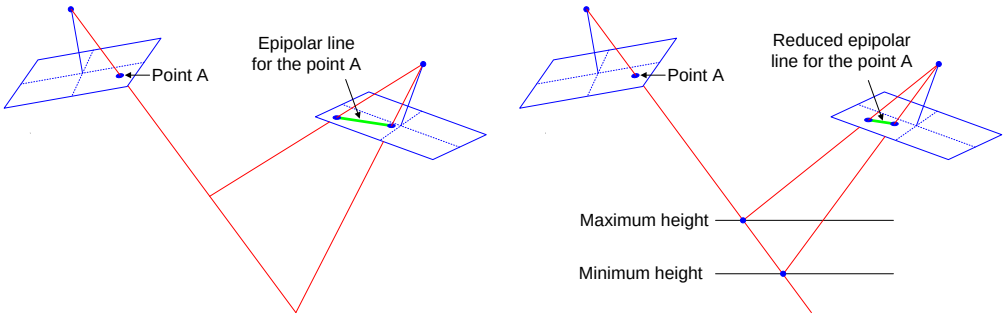


Figure 9. Conceptual diagram of epipolar line constraint (left) and vertical line locus constraint (right)

2.2.4 Space intersection

The space intersection calculates 3D coordinates of points that lie in the stereo overlap area. Figure 10 shows geometry of the space intersection.

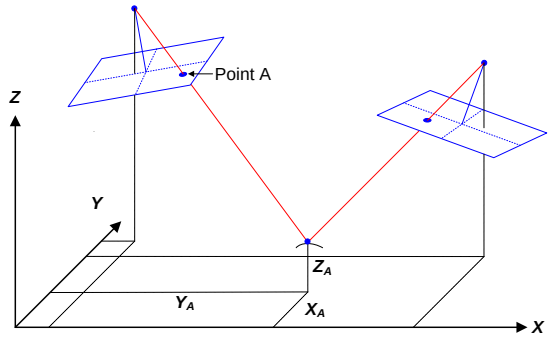


Figure 10. Geometry of the space intersection

2.2.5 Problems

Some objects (such as cylindrical roofs of the Air Force Museum buildings) show totally different textures when looking angles are different (Figure 11). Image matching performance is dramatically degraded in these areas.

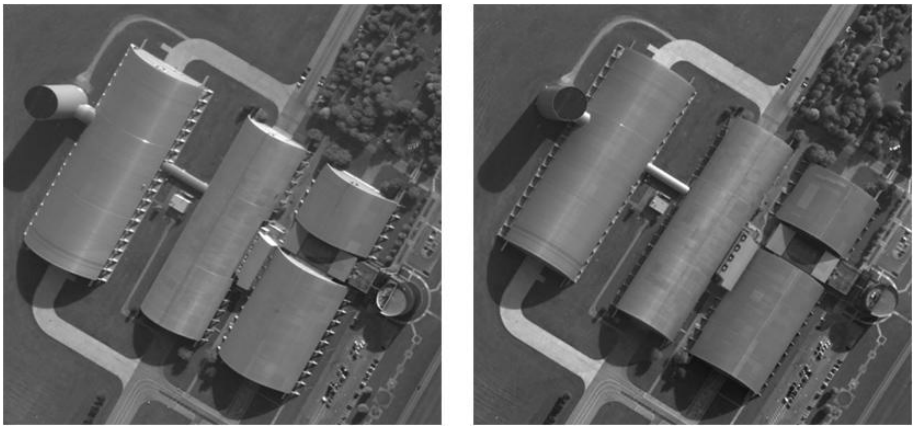


Figure 11. Cylindrical roofs of the Air Force Museum buildings

Lawn areas around airstrips (Figure 12) show low contrast and repetitive patterns which also degrade image matching performance.

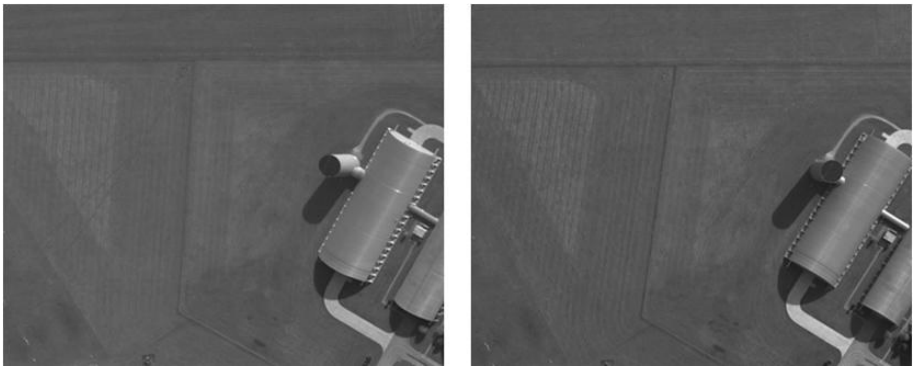


Figure 12. Lawn area around airstrips

2.3 Results

The result of image matching is 3D point cloud. Horizontal coordinates of the point cloud are in the UTM coordinate system; while heights are WGS84 ellipsoidal height. However, we can provide the results in any coordinate system. We generate surface using a set of MATLAB functions. Regular grid interpolation is needed to generate DSM (Digital Surface Model). This work will be done soon.

2.3.1 Wright-Patterson AFB

Figure 13 illustrates surface model of Wright-Patterson AFB generated from resulting 3D point cloud. In these surface models, red color represents highest surface while blue color represents lowest surface.

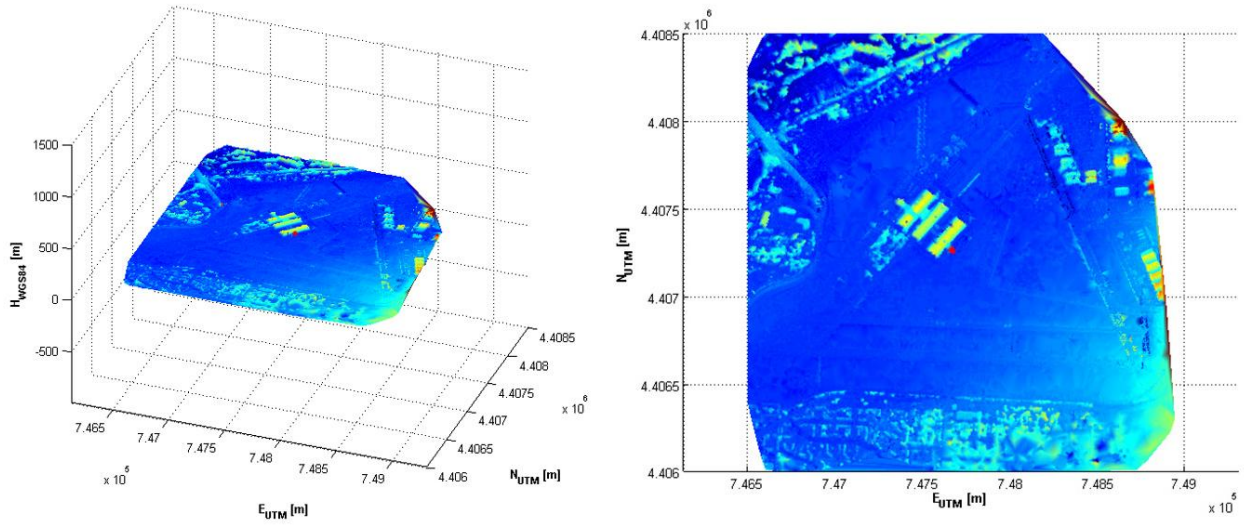


Figure 13. Surface model of the Wright-Patterson AFB

2.3.2 Five hangers

Figure 14 shows five hangers in south-west side of airstrips. Left figure shows surface model. Right figure shows point cloud projected on image.

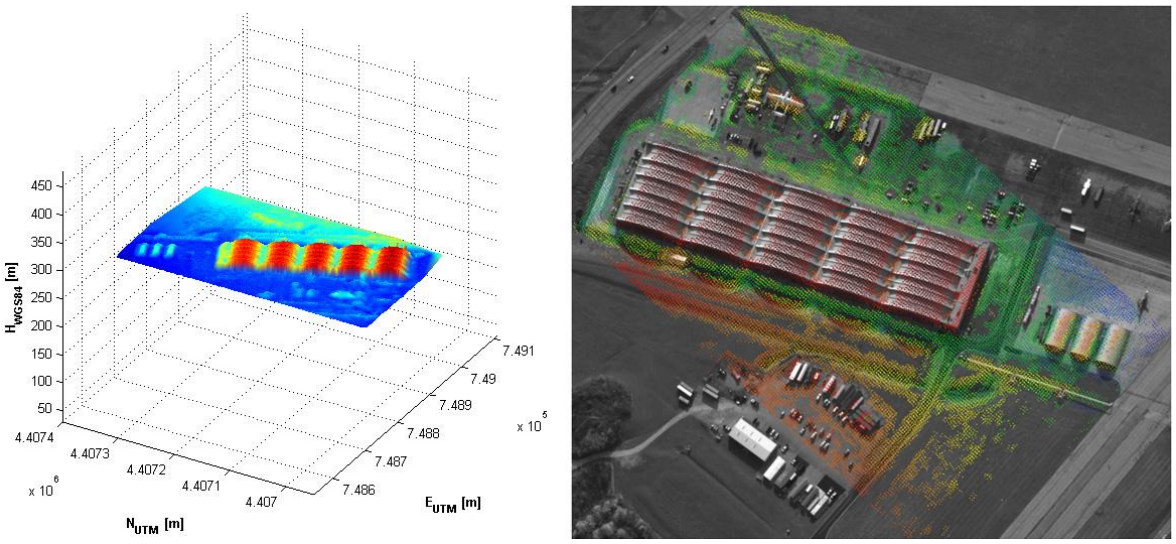


Figure 14. Surface model (left) and 3D point cloud projected on image (right) of five hangers

2.3.3 Air force museum

Figure 15 shows surface model and point cloud of the Air Force Museum. Problem in image matching of the cylindrical roofs is mentioned in 0.

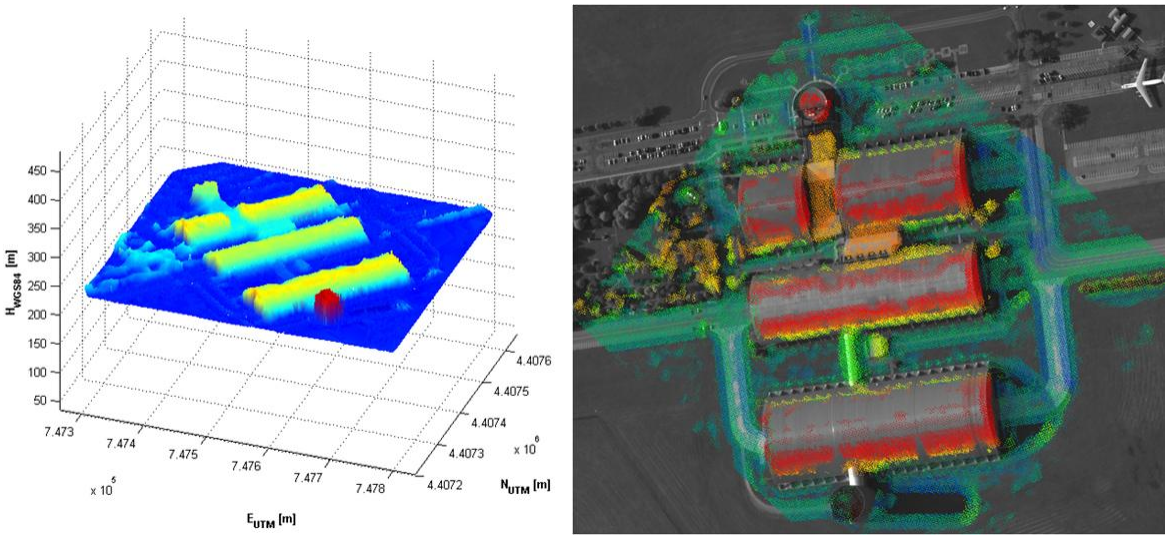


Figure 15. Surface model (left) and 3D point cloud projected on image (right) of Air Force Museum

2.3.4 Software

We developed *Mosaic Image Generator* and *DSM Generator*. Programs are developed in the Microsoft Visual C++ 6.0 environment. Current version of the Mosaic Image Generator is 1.31 and that of the DSM Generator is 1.11. We keep upgrading these programs. Figure 16 illustrates GUIs of these programs.

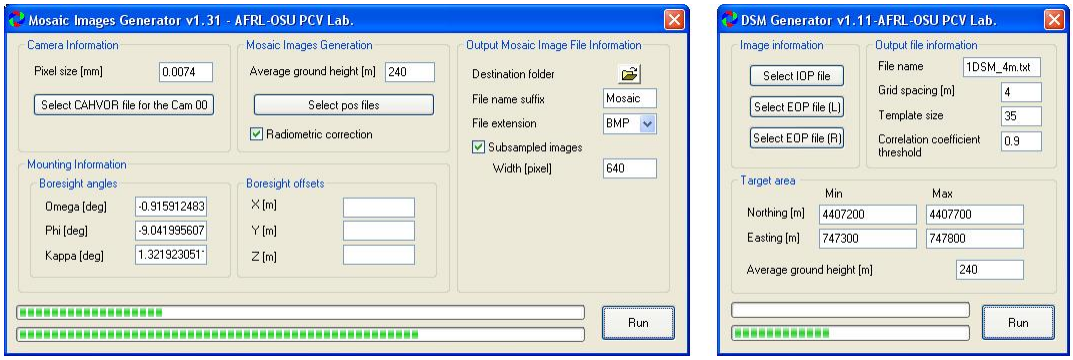


Figure 16. GUIs of developed programs

3. Evolving Point-Cloud Features For Gender Classification

Advances in sensor technology are driving demand for the development of new techniques for classifying 3D shapes. The key problem is finding salient features that can be quickly extracted from a sample of 3D point cloud data to form a signature suitable for use in a gender classification algorithm. Finding solutions to realistic classification problems (such as with 3D LIDAR) is often complicated by point cloud resolution and limited fields of view. Often sensors have fewer than a thousand points covering a restricted field of view. In this investigation, we utilize point clouds which cover the entire body (wrap around and head to toe). In order to establish a base line for more advanced research efforts, we bypass complications due to limited coverage and focused exclusively on achievable accuracy for point resolutions varying over three orders of magnitude with a minimum resolution of 100 points. Another major complication is the infinite number of possible articulations and orientations of the human body. Our data set is restricted to a finite number of standard poses, all of which have a definable vertical axis. Because we use full body coverage, the vertical axis is easily established using PCA for direction and the center of mass for location. Therefore, shape histograms based on point counts in concentric cylinders and/or cylindrical slices provide a natural basis for feature space representations. In this paper we derive these histograms from a finite number of concentric cylinders which are further divided with horizontal slices. Multilayer cylinders with slices and wedges are used to generate shape histograms for gesture recognition. Similar shape histograms generated with concentric spherical shells about a center of mass and additional sector models are used for shape similarity searches of 3D solids, for 3D shape matching and for human pose recognition.

A concentric cylinder is defined by three parameters specifying a radius and two positions on the vertical axis. Cylindrical histograms, which are conveniently defined by a set of parameter triplets, provide a very flexible ensemble for assembling effective feature vectors for gender classification. One can manually explore a small number of cylindrical histograms or employ soft evolutionary computing to automatically search for more optimal histograms. In this paper we explore the degree of improvement obtained with a conventional genetic algorithm using binary chromosomes that selects a subset of cylinders from a large predefined and fixed set of cylinders. Not investigated here, is the more general class of genetic algorithms which employs real valued chromosomes capable of representing parameter triplets and thereby capable of searching the entire space of possible histograms. However, the results of our preliminary investigation using binary selection chromosomes demonstrates that evolutionary computing is effective and necessary for the design of advanced point cloud classifiers.

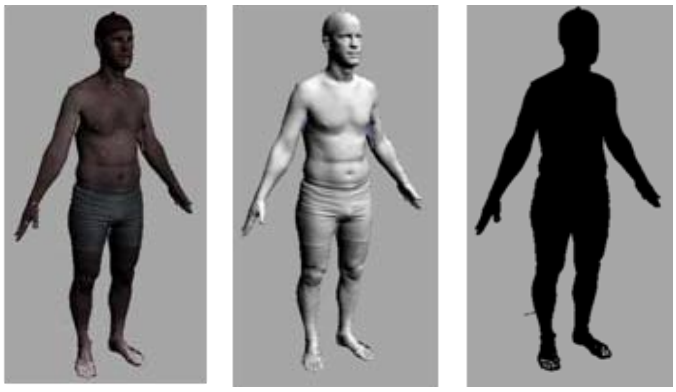


Figure 17: CAESAR Data. The leftmost image is a color polygon rendering of a subject using 316,691 polygon faces and 161,951 points. The small white dots on the surface of the subject are landmark points. The middle image is a grayscale rendering of the polygons. The rightmost image is the point cloud.

To test our system, we used data drawn from the CAESAR anthropometric database provided by the Air Force Research Laboratory (AFRL) Human Effectiveness Directorate and SAE International. A sample of the data available in the CAESAR database is shown in Figure 17. The database contains point clouds, mesh models and two types of measurements taken on approximately 4,400 human subjects. One group of measurements was taken by human experts using tape measures, calipers and scales while a second group of measurements was extracted from high resolution 3D LIDAR whole body scans of subjects wearing carefully placed physical markers that facilitate the automated extraction of important landmark locations. Both of these sets of measurements are carefully chosen based on extensive research in the area of anthropometric analysis and would be difficult to obtain using a sensor system in an uncontrolled environment. The traditional measurements require physical contact with the human subject while measurements dependent on landmark locations require the development of techniques for locating landmarks without the aid of physical markers on the human subject. Although the algorithmic identification of landmark locations is feasible, it is dependent on a relatively high resolution sensor scan and the ability to accurately locate specific points on a human subject many of which may be occluded in real world applications. Several techniques have been applied to solve the gender recognition using the traditional and extracted anthropometric measures. These techniques produce gender recognition accuracies above 98%.

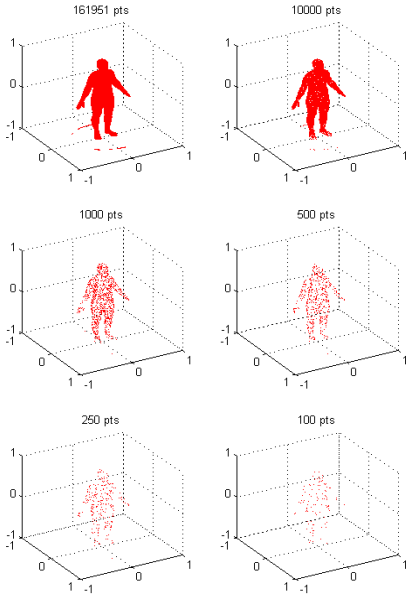


Figure 18: Point cloud resolution.

In this work, we use the raw 3D point clouds. The full resolution point clouds typically consist of 100,000 – 200,000 points. This resolution is ideal for developing meshes and analyzing various surface properties, but our focus is on the effect of taking point clouds from the CAESAR database and reducing the point density as much as three orders of magnitude as shown in Figure 18. Clearly the reduction from greater than 100,000 points to 1,000 points leaves sufficient structure to identify the cloud as a human shape. Even further reduction to 100 points maintains enough information to suggest a human form, but the question is can we discriminate gender using a low resolution point cloud and can we compensate for the loss of resolution by adapting the parameters controlling the shape histogram features?

3.1 Approach

Gender classification is performed using a traditional pattern recognition system consisting of a feature extraction module and a classifier module. The recognition system is embedded in a closed-loop evolutionary learning system that uses classification accuracy to evaluate the performance of different combinations of features. The evolutionary learning system varies the parameters of the feature extraction module to optimize the recognition accuracy. An overview of this system is shown in Figure 19.

3.1.1 Feature Extraction

To begin feature extraction, each point cloud is translated so the center of mass of the cloud is positioned at the origin of a 3D Cartesian coordinate system (X,Y,Z) with axes ranging from -1 to +1. The principal components of the cloud are computed and used to rotate the cloud so the largest principal component is aligned with the Z axis, the second largest component is aligned with the X axis and third largest component aligns with the Y

axis. The effect of applying these operations to a point cloud is shown to the left in Figure 20. A series of nested cylinders are superimposed over the aligned cloud such that long axis of the cylinder coincides with the Z-axis as shown to the right in Figure 4. Each cylinder is further subdivided into a series of slices. The total number of points in each cylinder slice is computed and serves as a rotationally invariant numeric feature with respect to the X-Y axes.

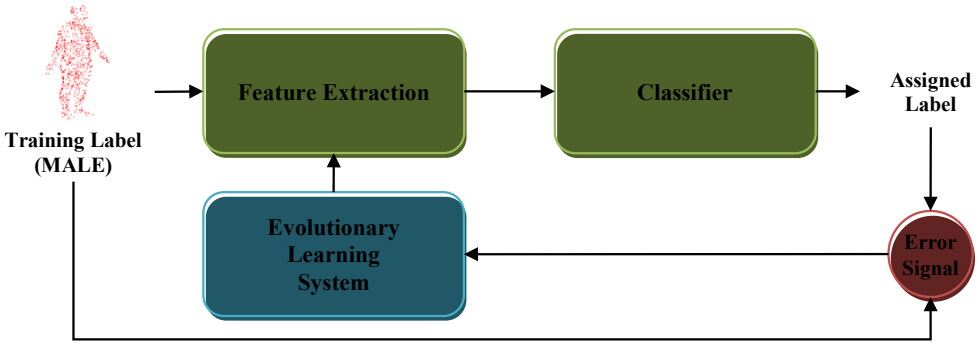


Figure 19: Closed loop evolutionary learning system.

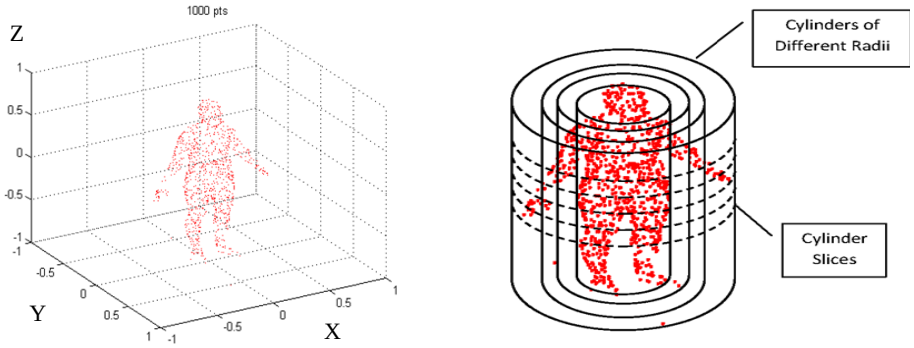


Figure 20: Cylindrical features.

The set of all numeric values derived from one cylinder forms a profile histogram as shown in Figure 21. The values for a set of cylinders define a set of profile histograms that captures a simplified view of a subject. To compensate for variations in the number of points in each histogram, each profile is normalized by the total number of points in the corresponding cylinder, converting the profiles into probability density functions. The concatenation of normalized components of density functions form a signature or feature vector that is passed to a classifier for labeling.

The sample histogram profiles shown in Figure 21 capture the essence of human subjects. Cylinders with small radii are sensitive to smaller physical features such as the head as seen in the rightmost portion of the topmost profile in the figure. The zero values in the topmost profile occur because the subject's torso exceeds the radius of the cylinder. Each histogram shows variations in different areas based on the subject's physique. Roughly, the small cylinders capture the features related to the head, midsize cylinders contain torso features and the largest cylinders hold everything else including outstretched limbs. We

should note that this representation is designed to measure human subjects in a normal urban environment where people are walking or standing in an upright position.

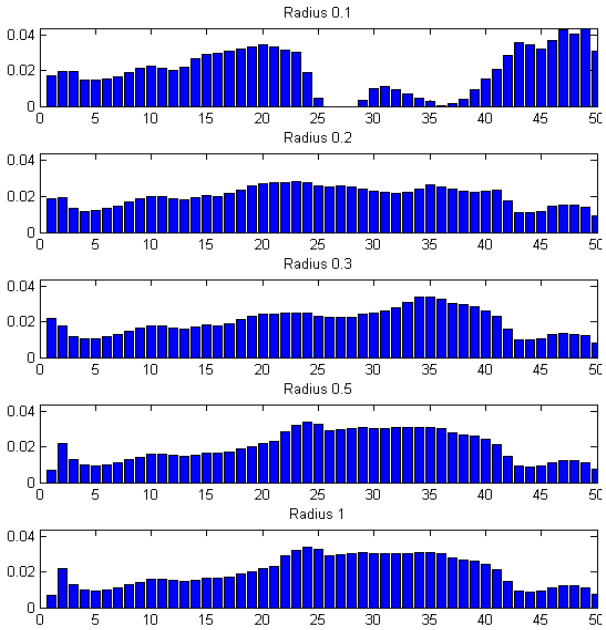


Figure 21: Histogram derived from cylindrical features.

Our choice of representation has several interesting properties. Once the point cloud is properly oriented, the features are rotationally invariant with respect to the X-Y plane. In addition, the cylinders are nested and the point counts of inner and outer cylinders are not mutually exclusive. For example, points contained in a cylinder of radius R are also counted in all other cylinders with radii greater than R. The use of nested cylinders creates an inherent redundancy in the representation that reduces the need to find a highly tuned set of cylinders with specific radii for a given data set.

3.1.2 Classifier

The classifier is implemented using a support vector machine (SVM) available in the WEKA machine learning software. The SVM is a classification technique suitable for solving two-class classification problems. There are many variations of support SVM each having attributes that allow the user to customize the SVM to the specific characteristics of a given classification problem. In general a SVM forms a model of a labeled input data set and fits a maximum- margin hyper plane between the two classes of data. The WEKA software uses a sequential minimal optimization technique to accelerate the process of training the classifier.

3.1.3 Evalutionary Learning System

We use a traditional genetic algorithm11 to optimize the set of cylinder radii and evaluate the performance of a given set of cylinders by extracting features from sample point cloud data to determine the gender classification accuracy. To apply a genetic algorithm, we need

to determine a representation of our search space, choose a performance evaluation function, select specific parameters of the evolutionary algorithm and choose a termination condition.

Cylinder Radius	0.05	0.10	0.15	0.20		0.85	0.90	0.95	1.00
Chromosome 1 (Best)	0	1	0	1	…	1	0	0	1
Chromosome 2	1	0	0	1	…	0	1	0	1
Chromosome 3	0	1	1	0	…	1	0	1	1
…									
Chromosome 99	1	0	0	0		0	1	0	1
Chromosome 100	0	1	0	1		1	0	1	1

Figure 22: Chromosome representation and sample population.

We chose to represent a potential solution as a chromosome composed of 20 bits as shown in Figure 22. Each bit represents whether or not a cylinder of a specific radius is included in the solution. For our experiments we allowed for 20 cylinders of radii ranging from 0.05 to 1.0 in increments of 0.05. We chose to always include the cylinder of radius 1.0 in the solution to ensure that all data points were measured in some feature. This representation defines a search space of 219 possible configurations of cylinders.

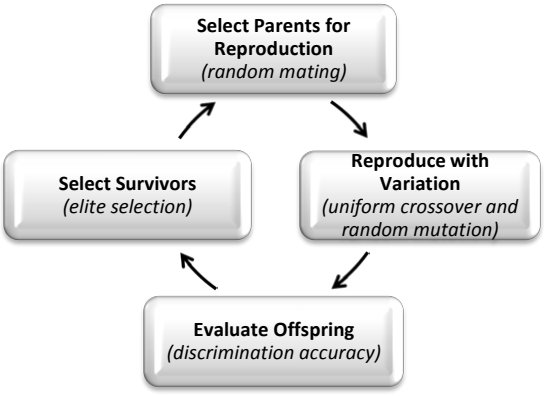


Figure 23: Evolutionary learning cycle.

To begin the evolutionary search, a population of 100 chromosomes was generated. Each bit was initialized to 0 or 1 at random. Each chromosome contained on average 10 cylinders of varying radii. The performance of a set of cylinders was evaluated by extracting the features from a sample of point cloud data and measuring the gender classification accuracy. This accuracy was used to score the fitness of the cylinder configuration. This process was repeated for each member of the population and the chromosomes were rank ordered by accuracy.

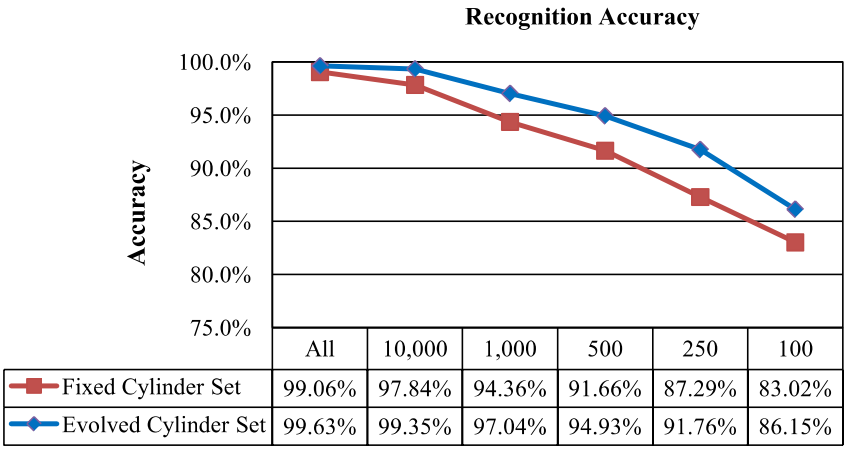
The evolutionary cycle is shown in Figure 23. Pairs of parental chromosomes are selected at random for mating. A new offspring is formed using uniform crossover12. Uniform crossover selects the value for each bit position in the offspring by randomly choosing the value of one of the parental bits in the corresponding position. Once the basic structure of the offspring is formed some bits are randomly mutated (0 to 1 or 1 to 0) to introduce

further variation into the offspring. Each offspring is evaluated by scoring the fitness of the cylinder configuration for gender classification accuracy. A steady- state genetic algorithm is used with a (μ+λ) elite selection strategy with μ = 100 and λ = 20. There is a penalty function incorporated in the selection process because solutions with equal accuracy are rank ordered by the number of cylinders. This induces a small selective pressure to evolve solutions with fewer features. To summarize, the population expands from 100 parental chromosomes to 120 chromosomes (parent + offspring) and is culled back to 100 individuals using elite selection with a complexity penalty before the next generation of the evolutionary cycle begins. This process is repeated for 100 generations.

In terms of the search space, we begin with an initial sample of 100 configurations and generate 2000 new configurations (100 generations x 20 offspring). This represents a total of 2100 sample configurations drawn from a search space of 219=524,288. Thus, the genetic algorithm explores approximately 0.4% of the total search space in an effort to find an improved cylinder configuration.

3.2 Results

To establish a baseline level of performance, we defined a fixed set of cylinders with specific radii based on a human expert's estimate of the scale of the most salient regions of the point clouds. Five cylinders of radii 0.1, 0.2, 0.3, 0.5 and 1.0 were selected to capture physical attributes of the head, torso and whole body shape. We then proceeded to explore the efficacy of using a variable number of cylinder features and the impact on classification accuracy as a function of data density. The data samples were divided evenly into a training set and a validation set. The training data was used to optimize the choice of sets of cylinder sizes at each point cloud density level while the validation set was sequestered for use at the end of the evolutionary process to test the accuracy of the final evolved solution.



	All	10,000	1,000	500	250	100
Fixed Cylinder Set	99.06%	97.84%	94.36%	91.66%	87.29%	83.02%
Evolved Cylinder Set	99.63%	99.35%	97.04%	94.93%	91.76%	86.15%

Figure 24: Classification Accuracy Using All Data For Training.

The results of this experiment are shown in Figure 24. The numerical classification accuracy is shown in the table below the figure. The fixed cylinder set refers to a set of cylinders of radius 0.1, 0.2, 0.3, 0.5 and 1.0 while the evolved set is the set of cylinders in the most accurate solution after 100 generations of the genetic algorithm. Each experiment was replicated three times to make sure the results were consistent, but only the results for a single replicate are reported in the figure. The chromosome corresponding to the best evolved solution for each point cloud density is shown in Figure 26. The recognition accuracy results are very good for both the fixed radii and the evolved radii cylinder sets, but clearly the evolved set is consistently superior. The classifier performance increase ranges from 0.5% to 3% with larger performance increases for cloud densities less than or equal to 1000 points. The maximum difference occurs for the 250-point density experiment where the evolved solution achieved a four percentage point increase in overall accuracy compared to the fixed radii cylinder solution. This represents approximately 175 additional samples being correctly classified. These results clearly suggest that even when a specific type of feature has been selected for a given problem (e.g., cylinder / histogram point counts), specific operating conditions present when the data is collected (e.g., sensor system characteristics), might be used to parameterize the feature extraction system to improve overall performance.

Points	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
100	0	0	1	1	1	1	1	1	0	1	0	1	1	1	0	0	0	0	1	1
250	1	0	1	1	1	1	1	0	0	1	1	1	0	0	0	1	0	0	0	1
500	1	1	1	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	1
1000	1	1	1	1	0	1	1	0	1	1	0	0	0	0	0	0	0	0	0	1
10000	1	1	1	1	1	1	0	0	1	0	0	0	1	1	0	0	0	0	0	1
All	1	1	1	1	1	1	0	0	1	0	0	0	1	1	0	0	0	0	0	1

Figure 25: Best evolved chromosomes.

A more detailed examination of the results shown in Figure 25 supports a conjecture that cylinders with large radii are not as useful for gender classification as cylinders with relatively small radii. Recall that the cylinder of size 1.0 is forced on, but no other larger radii are selected. All solutions contain cylinders of radii 0.15, 0.20 and 0.35. These sizes roughly correspond to the fixed size cylinders of radii 0.1, 0.2 and 0.3 that capture head and torso features. The results suggest that the evolutionary algorithm is able to fine tune these radii to produce a more accurate result. One additional observation is there is a tendency to favor cylinders with small radii when the resolution of the point cloud is high. For example, the average radius of all cylinders used for the 100 point density cloud is 0.30 while the average cylinder size for the 10,000 point density cloud is 0.19. This may indicate that when point densities are high, the quality of the smaller features found in the head region are more reliable, but additional experiments are needed to confirm this observation.

Figure 26 provides insight into the evolutionary process and the ultimate choice of radii for a dataset of a given resolution for a given resolution of data set. Each plot in this figure measures the probability of 0 or 1 in the population as a function of generation. The color

in this figure represents the probability of a 1 in a specific bit position. A dark blue color indicates a near zero probability of a 1 in the given bit position while a dark red color indicates the probability of a 1 in a bit position is approaching 1.0. The range of colors and their associated probabilities are shown in the small bar to the right of each plot. The first couple of lines in each plot are a mid-range color because the distribution of bits in the population are random so the probability of a 1 bit is approximately 0.5. As the evolutionary process cycles through generations the distribution change, but the rate at which certain bit positions converge is quite different. The bit positions associated with smaller radii tend to converge most quickly. The bit corresponding to the cylinder with radius 2.0 appears to converge to 1 within 5 generations regardless of the resolution of the data. Similarly the cylinder with radius 0.15 converges within 10 generations. Cylinders with these two radii These two radii contain the head of the subjects. Bit 6 that represents a cylinder of radius 3.0 also is present in all resolutions of data. A cylinder of radius 3.0 would capture the torso. Recall that the cylinder of radius of 1.0 is intentionally included so it is not part of the search. We can also see an interesting behavior in the first bit position. It appears that this small cylinder size is useful solutions with point cloud resolutions of 250, 500, 1000 and 10000 points, but not useful for resolutions of either 100 or ALL (>100,000 points). This seems like an anomaly, but in fact it is entirely consistent with the strong selective pressure induced by elite selection for survival. We would expect every bit position in the population to eventually become homogeneous (e.g., every individual in the population has 0 in the position or every individual has a 1 in the position). Bit positions that quickly converge to 1 regardless of the resolution of the data may indicate high discriminatory value. Bit positions that slowly converge to zero are being explored and rejected. The same pattern that was observed as depicted in Figure 9 is visible in this series of plots. There appears to be a subtle tendency to use a mixture of smaller radii cylinders for high resolution samples and a more diverse range of slightly larger cylinder sizes for low resolution samples.

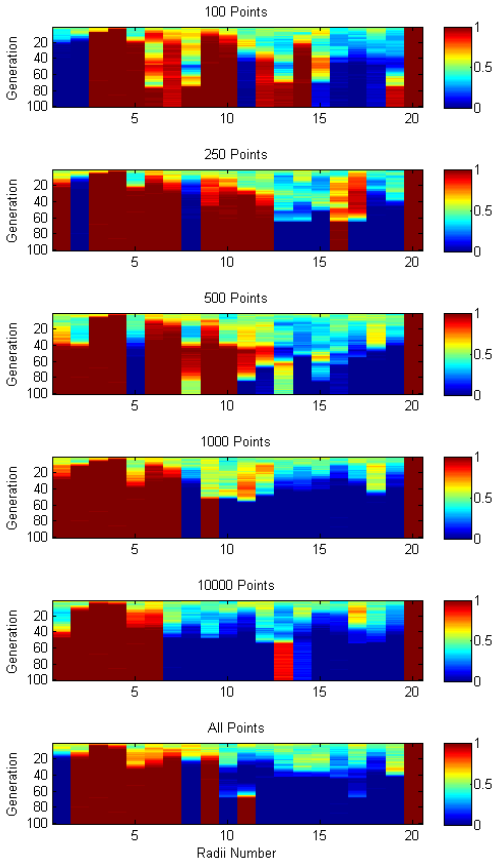


Figure 26: distribution of cylinder radii in evolved populations.

3.3 Discussion

We present a preliminary program to investigate the application of evolutionary computing to the design of gender classification classifiers using point cloud data. This investigation used the CAESAR point cloud database which ideally provides complete body coverage with high point density. The full coverage allows one to identify a vertical body axis so that concentric cylindrical regions are definable and useable for generating feature vectors from cylindrical histograms. Surprisingly, this type of feature vector was found to be effective as well as efficient for gender classification using point clouds. Therefore the full coverage feature was maintained while the number of points per sample was varied over three orders of magnitude. While for most applications the availability of such ideal coverage is not realistic, our results provide an important baseline for further development.

Combining cylindrical histograms with SVM-based discriminators yields impressive gender recognition results. With high resolution point clouds the evolved recognition system achieved 99.6% accuracy with approximately 4400 samples. Figure 26 summarizes the dependency of classification accuracy on the number of points. When point cloud density is reduced by one order of magnitude (100K to 10K) the algorithm's accuracy in distinguishing between genders remains at 99.3%. When the density is reduced by two orders of magnitude (100K to 1K) accuracy is still preserved at 97%. Significant

degradation in performance does occur when the density is reduced by three orders of magnitude, but an accuracy of 86% is still observed. At this resolution, a human expert would have difficulty consistently distinguishing gender.

As discussed above, the cylindrical histograms are well suited to an evolutionary search process. A simple genetic algorithm with binary chromosomes for selecting optimal subsets of cylinders demonstrated significant improvements over all data cloud point densities. As depicted in Figure 26, the improvement over the fixed cylinder set increased from one percentage point at the highest density to four percentage points at the lowest densities. The final recognition rates rivaled those achieved using hand-measured anthropometric features even when the density of points on target was relatively low. An interesting observation is that the choice of features varied with the density of the point cloud. As seen in Figure 11, coarse feature measurements (larger radii) are more effective for gender classification using low resolution point clouds while fine grained features were more effective for high resolution point clouds. This result suggests that even when a specific type of feature is used for a given application, adapting some aspect of the features to compensate for variations in sensor measurements, can produce a significant increase in performance. Such properties are critical to the development of the next generation of robust security systems.

4. Future works

4.1 Regular grid interpolation of resulting 3D point cloud

Eventually, 3D point clouds should be interpolated to generate regular grid DSM (Digital Surface Model). Several interpolation methods can be applied. This work will be done soon.

4.2 Compensating imperfect time synchronization of imaging

Imperfect time synchronization makes synthetic images inaccurate. For example, if a camera takes image 1/1000 second later than intended timing, the position of the perspective center will show 3 inches difference, which is huge, from intended position when airspeed is 180 mph. Also, there should be platform deformation that slightly changes positions and angles between camera heads due to highly dynamic environment of the flight mission. All commercial multi-head camera system provider use image matching and bundle adjustment using images acquired from same time epoch to minimize errors due to imperfect time synchronization as well as platform deformation.

4.3 Estimating lens distortion parameters

Lens distortion parameters are very important to generate accurate 3D models. However, given CAHVOR model does not have lens distortion parameters. We can calculate rough lens distortion parameters with given images. However, lack of ground control points and obliqueness of images only allow calculating rough lens distortion parameters. We can calculate more accurate lens distortion parameters with several new images of our calibration panel.

4.4 Direct geo-referencing

Given navigation solution (pos files) cannot be used directly; because, there are boresight misalignment angles (unknown angles between navigation system and camera system) as well as offset vector (positional displacements between navigation system and camera system). We estimated only boresight misalignment angles by using bundle adjustment with a number of GCPs (Ground Control Point). The exterior orientation parameters directly calculated from pos files and the boresight angles are not acceptable for surface modeling (very close, but not acceptable; there are less than two meters gaps on the ground between estimated values from bundle adjustment and calculated values from pos files). At this time, we are not sure whether given navigation solutions provide acceptable accuracy or not. Also, there could be unknown offset vector. We will estimate the vector first; then we will determine acceptability of the given navigation solution. However, unknown quantities such as lens distortion parameters could disturb accurate estimation of parameters. It will be helpful if the specification of the navigation system is given.

4.5 Adaptive template matching

Enhancing matching performance (an area for future work).

5. Remarks for further flight missions

5.1 Overlaps between images

Overlapping areas between images are important for radiometric correction and compensating imperfect time synchronization. Given raw images show that brightness and contrast are different even images are acquired at same time. To generate seamless mosaic image, images should have overlapping areas. Compensating imperfect time synchronization is mentioned in 4.2.

5.2 Lens distortion parameters

The importance of the lens distortion parameter is described in 4.3.

5.3 Focal length

The focal lengths of the all six physical cameras should be fixed and precisely measured before and after mission for further photogrammetric process. For the aerial camera, focusing mode does not need to be auto-focusing; because, object distance is extremely larger than image distance (distance from lens center to image plane); if focal length is same with image distance, any object on ground is always clearly focused.

Appendix (Analygraph)

Wright-Patterson AFB

The Air Force Museum

Five hangers

AFRL buildings